REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
	New Reprint	-

4. TITLE AND SUBTITLE
Perceptions of randomized security schedules.

5a. CONTRACT NUMBER
W911NF-11-1-0332

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
611103

6. AUTHORS
Scurich, N., John, R. S.

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES
University of Southern California
3720 S. Flower Street

Los Angeles, CA 90089 -0701

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S)
ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
59733-NS-MUR.110

12. DISTRIBUTION AVAILIBILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT
Security of infrastructure is a major concern. Traditional security schedules are unable to provide omnipresent coverage; consequently, adversaries can exploit predictable vulnerabilities to their advantage. Randomized security schedules, which randomly deploy security measures, overcome these limitations, but public perceptions of such schedules have not been examined. In this experiment, participants were asked to make a choice between attending a venue that employed a traditional (i.e., search everyone) or a random (i.e., a probability of being searched) security schedule. The absolute probability of detecting contraband was manipulated (i.e., 1/10, 1/4, 1/2) but

15. SUBJECT TERMS
Judgment and decision making; legal policy; security; terrorism

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Milind Tambe
UU	UU	UU	UU		19b. TELEPHONE NUMBER 213-740-6447

Report Title

Perceptions of randomized security schedules.

ABSTRACT

Security of infrastructure is a major concern. Traditional security schedules are unable to provide omnipresent coverage; consequently, adversaries can exploit predictable vulnerabilities to their advantage. Randomized security schedules, which randomly deploy security measures, overcome these limitations, but public perceptions of such schedules have not been examined. In this experiment, participants were asked to make a choice between attending a venue that employed a traditional (i.e., search everyone) or a random (i.e., a probability of being searched) security schedule. The absolute probability of detecting contraband was manipulated (i.e., 1/10, 1/4, 1/2) but equivalent between the two schedule types. In general, participants were indifferent to either security schedule, regardless of the probability of detection. The randomized schedule was deemed more convenient, but the traditional schedule was considered fairer and safer. There were no differences between traditional and random schedule in terms of perceived effectiveness or deterrence. Policy implications for the implementation and utilization of randomized schedules are discussed.

REPORT DOCUMENTATION PAGE (SF298)
(Continuation Sheet)

Continuation for Block 13

ARO Report Number 59733.110-NS-MUR
Perceptions of randomized security schedules. ...

Block 13: Supplementary Note

Approved for public release; distribution is unlimited.

Risk Analysis, Vol. 34, No. 4, 2014 DOI: 10.1111/risa.12126

Perceptions of Randomized Security Schedules

Nicholas Scurich[1,2,*] and Richard S. John[3,4]

Security of infrastructure is a major concern. Traditional security schedules are unable to provide omnipresent coverage; consequently, adversaries can exploit predictable vulnerabilities to their advantage. *Randomized security schedules*, which randomly deploy security measures, overcome these limitations, but public perceptions of such schedules have not been examined. In this experiment, participants were asked to make a choice between attending a venue that employed a traditional (i.e., search everyone) or a random (i.e., a probability of being searched) security schedule. The absolute probability of detecting contraband was manipulated (i.e., 1/10, 1/4, 1/2) but equivalent between the two schedule types. In general, participants were indifferent to either security schedule, regardless of the probability of detection. The randomized schedule was deemed more convenient, but the traditional schedule was considered fairer and safer. There were no differences between traditional and random schedule in terms of perceived effectiveness or deterrence. Policy implications for the implementation and utilization of randomized schedules are discussed.

KEY WORDS: Judgment and decision making; legal policy; security; terrorism

1. INTRODUCTION

The threat of terrorism constitutes a major societal concern.[1,2] The attacks on September 11, 2001 invigorated efforts to protect national infrastructure, including historical landmarks, places of political or economic importance, power generation facilities, and airports.[3] Traditional wisdom is that security measures ought to be omnipresent. At an airport, for example, this would entail monitoring all entrances, inbound traffic, and persons. However, limited resources realistically preclude this possibility. As a result, adversaries can observe security arrangements over time, and exploit any predictable vulnerabilities to their advantage.

One way to countervail this type of exploitation is to use a *randomized security schedule*. The Los Angeles International Airport (LAX) has implemented such an approach, which uses a computer program to create a schedule that determines what, where, and when certain security measures will be deployed.[4] For example, police canine units might be deployed to terminals 1, 3, and 5 on one day, and terminals 2, 4, and 6 the next day. Similarly, vehicular checkpoints might be implemented between 8 a.m. and 10 a.m. on one day and 1 p.m. and 4 p.m. the next day. This same approach could be used to select passengers to be subjected to heightened levels of scrutiny and search.[4] The unpredictability of the schedule makes it exceedingly difficult for adversaries to discern an exploitable pattern.[5]

Randomization algorithms are based on a game-theoretic solution to a Bayesian Stackelberg game. In brief, a Stackelberg game allows one to model interactions between a defender (e.g., security forces) and

[1]Department of Psychology and Social Behavior, University of California, Irvine, CA, USA.
[2]Department of Criminology, Law and Society, University of California, Irvine, CA, USA.
[3]Department of Psychology, University of Southern California, Los Angeles, CA, USA.
[4]Center for Risk and Economic Analysis of Terrorism Events (CREATE), USC.
[*]Address correspondence to Nicholas Scurich, Department of Psychology and Social Behavior, 4312 Social & Behavioral Sciences Gateway, Irvine, CA 92697-7050, USA; nscurich@uci.edu.

an attacker (e.g., an adversary), both of whom (1) pursue self-interests and (2) are intelligent players. The attacker (follower) is assumed to have perfect knowledge of the defender (leader) strategy, as defined by the probability distribution that the defender will guard each of the possible targets. Furthermore, both the attacker and defender are assumed to have perfect knowledge of the payoff matrix of utilities to both sides for both successful and unsuccessful attacks on any given target. In general, the payoff matrix is not assumed to be zero sum. The Stackelberg solution dictates the optimal course of action for the defender and the attacker in the sense that any deviation will result in a lower expected value for the side that defects. In the present context, the key aspect of the algorithm is to specify the probability distribution for guarding particular targets that maximizes the expected utility of the defender. In the standard formulation, the attacker adversary is assumed to be a rational agent wanting to maximize expected utility.

Although there are theoretical advantages to randomized security strategies, and they have been implemented in several major areas, there has been no research evaluating the public's perception of such measures. Public perceptions are relevant for several reasons. First, the public is the primary stakeholder group that drives public policy formulation. Even sound, scientifically informed policy will be eschewed if the public detests it. Second, and perhaps more importantly, the public must believe that the measures do in fact promote security. Mistrust of security could cause potential passengers to utilize alternative forms of transportation, which might actually be more risky. For example, Gigerenzer[6,7] demonstrated that the fear of flying caused by the attacks on September 11, 2001 led people to drive instead of fly, and that the number of driving fatalities significantly increased as a result. Similarly, Su *et al.*[8] found evidence for a proximity effect, in which traffic fatalities increased significantly only in the Northeast during the three months following the 9/11 attack, as well as a significant increase in fatal accidents involving alcohol or drugs in the Northeast during the last three months of 2001. And a recent follow-up analysis by Gaissmaier and Gigerenzer[9] confirmed the proximity effect by demonstrating that both driving and fatalities increased inversely as a function of distance from New York City.

Perhaps the most challenging hurdle for randomized security strategies is the potential for perceived unfairness by the public.[10] Randomization is clumpy,[11] and random selections for search will often appear nonrandom to an individual who observes some relatively small number of searches while waiting in line. Members of the public are likely to apply the "law of small numbers,"[12] expecting to see perfectly even (or representative) distributions of searches for all genders, ages, ethnicities, and nationalities. Such misperceptions of randomness[13,14] are likely to lead many members of the public to believe that truly random search schedules are in fact nonrandom and biased toward selection of individuals from particular groups. The tendency to see patterns in small samples may also be exacerbated by suspicions of inequity and unfairness, prior to any direct observation of searches. Individuals observing random searches may exhibit a confirmation bias[15] that magnifies the perception of unfair search patterns in short sequences due to (1) biased observation seeking to confirm hypothesized inequities and (2) biased interpretation and recollection of observed searches.

In short, the perception of safety can be as important as the reality of safety. Similarly, the perception of fairness can weigh as heavily as the reality of fairness. If randomized security schedules are perceived as inefficacious and/or unfair, potential patrons might protest their use and pursue alternatives that actually increase the net societal risk. The present experiment provides a test of the relevant perceptions.

2. METHODS

Participants were adult U.S. residents (age 18 years and older) who completed an online survey posted on Amazon Mechanical Turk (AMT). Several studies have investigated the representativeness of AMT samples, and found them to be superior to typical convenience samples obtained locally.[16–19] Each participant received a nominal payment for his or her effort. After removing those who failed a memory-check question ($n = 50$),[20] the sample consisted of 214 participants (median age = 31 years; IQR = 26–40 years), of which 55% participants were male.

In brief, the task elicited a choice between two competing security search options. For instance, participants were told:

> Assume that you must fly across the country for an important meeting. There are two airports that are equidistant from your residence. The airports are identical in every respect, but differ only in the security screening procedure they use.

> Airport A has a procedure that searches all passengers. It will detect 1 in X passengers who carry contraband (*traditional approach*).
>
> Airport B has a procedure that randomly searches 1 in X passengers. Of the passengers selected to be searched, it will detect all who are carrying contraband (*randomized approach*).[5]

Participants were asked which option they would choose and rated their strength of preference on a seven-point Likert scale. Note that the use of a forced choice better matches decisions people actually make when choosing between two available options, rather than independently evaluating alternatives in the abstract. Further note that frequencies were utilized in order to increase the comprehensibility of the numerical information provided to participants.[21,22]

On a separate screen, participants then compared the two options along the following five dimensions: overall safety, fairness of the security measure, effectiveness of the security measure, convenience of the security measure, and likelihood to deter contraband. These ratings were made on a seven-point bipolar scale with the competing options on opposite ends and a middle point (= 4) indicating indifference.

There were three different probabilities of detection (i.e., 1/10, 1/4, and 1/2) in this repeated-measures design. The probability of detection was equivalent between the two options (traditional vs. random approach). That is, the traditional approach searched everyone but had a 1 in X chance of detecting contraband for every person, while the randomized approach searched 1 in X people but had a 100% chance of detecting contraband for those who are searched. In order to use a within-participants design, we created two other security contexts: a stadium and a vehicular checkpoint. Each of the three probabilities was paired with a different security context in each group, and each probability appeared in a different position (first, second, or third) in each of the three groups. As the contexts were not of substantive interest, the ordering and context pairing of the probability conditions was counterbalanced, while the contexts always appeared in the same order. For example, $p = 1/10$ appeared in the first position (airport) for group 1, in the second position (stadium) for group 2, and in the third position (vehicular checkpoint) for group 3. Participants were assigned randomly to one of three groups. Thus, all participants experienced three different contexts and three different probabilities.

[5]The order in which the traditional or random options appeared was counterbalanced.

3. RESULTS

Participants were about evenly split regarding preference for the randomized versus traditional approach; slightly more than half (51.9%, 51.4%, and 51.9%) of the participants preferred the randomized procedure in the three conditions (p(detection) = 1/10, 1/4, and 1/2), respectively. There were no significant differences among the three conditions regarding proportion preferring randomized versus traditional search, and none of the proportions were significantly different from 0.50.

All ratings of security procedures were transformed from a 1 to 7 scale to a –3 to +3 scale, with 0 indicating equivalence between the random and traditional approaches. Positive values indicate a preference for the randomized approach and negative values signify a preference for the traditional approach. Mean scores with 95% confidence intervals (± 2 *SE*) are displayed in Fig. 1 for all five attributes under each of the three search conditions.

For each of the five attributes rated, a multivariate analysis of variance (MANOVA) was conducted for all three search conditions to determine whether the means (intercept term) are different from zero, indicating a preference for either the randomized or traditional approach. The traditional approach was rated as both safer and fairer than the randomized approach (means less than 0.0), $F(3,211) = 4.03$, $p < 0.01$, $\eta_p^2 = 0.054$ and $F(3,211) = 27.43$, $p < 0.001$, $\eta_p^2 = 0.281$, respectively. Univariate tests indicate that subjects rated the traditional approach as significantly safer in the $p = 1/4$ condition only, with no significant differences from 0.0 in either the $p = 1/10$ or $p = 1/2$ conditions. Univariate tests indicate that the traditional approach was rated as significantly fairer in all three search conditions. The randomized approach was rated as more convenient than the traditional approach (means greater than 0.0), $F(3,211) = 25.67$, $p < 0.001$, $\eta_p^2 = 0.267$. Univariate tests indicate that the random search strategy was perceived as more convenient in all three conditions. Participants rated the traditional and random search procedures as nearly equal with respect to both effectiveness ($F(3,211) < 1.0$) and deterrence ($F(3,211) = 1.19$, n.s.).

For each of the five attributes, a $3 \times 2 \times 2$ mixed model MANOVA was also conducted to test for differences by search condition ($p = 1/10$, 1/4, and 1/2),

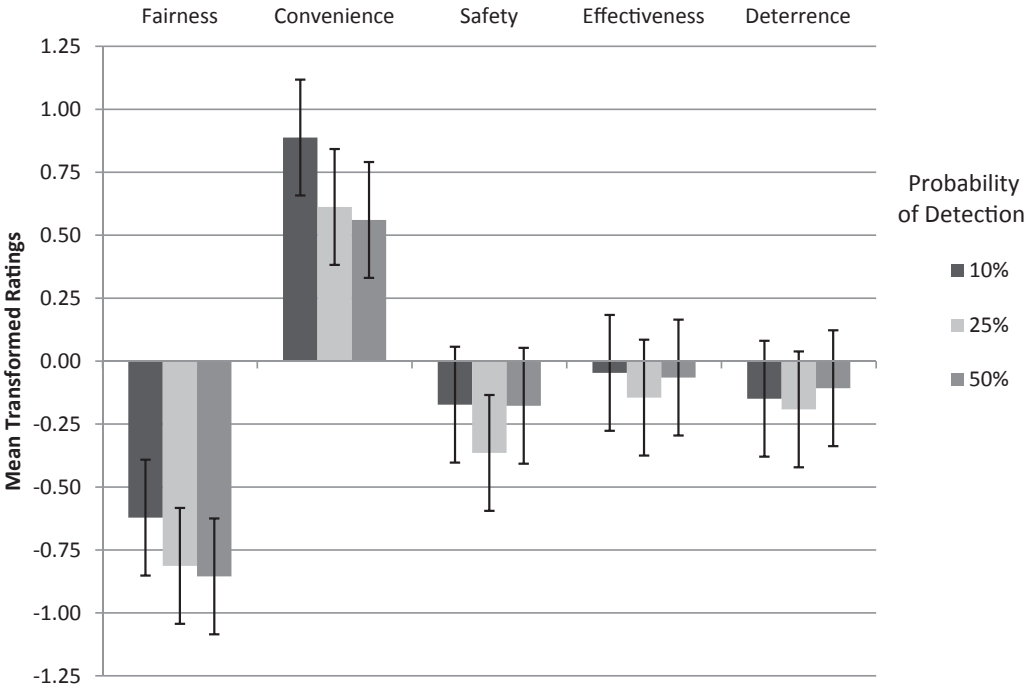

Fig. 1. Relative preference for randomized versus traditional security schedules. Note that positive values indicate a preference for a randomized security schedule and negative values indicate a preference for a traditional security schedule.

sex (male vs. female), and age (18–31 vs. over 31). Subjects' perceptions of the relative safety, fairness, effectiveness, and deterrence of the randomized versus traditional strategies were remarkably consistent across the three search conditions ($p = 1/10$, 1/4, and 1/2). The search condition's main effect for each of the four attributes was not significant. Furthermore, relative ratings of safety, fairness, effectiveness, and deterrence did not differ by either sex or age; all of the main effects for these two subject variables, and the sex by age interaction, were not significant. Neither sex nor age served to moderate the effect of search condition on relative ratings of safety, fairness, effectiveness, or deterrence; all interactions involving search condition with sex or age or both were not significant for these four attributes.

Participants did rate the three search conditions significantly differently in terms of relative convenience of the randomized versus traditional strategies, $F(2,209) = 3.127$, $p = 0.046$, $\eta_p^2 = 0.029$. The linear contrast for search condition was significant, $F(1, 210) = 5.701$, $p = 0.018$, $\eta_p^2 = 0.026$; the lower the proportion of people stopped and searched in the random search condition, the greater the advantage of the randomized over the traditional approach. In addition, the age by search condition interaction for relative convenience ratings was also significant, $F(2, 209) = 3.785$, $p = 0.024$, $\eta_p^2 = 0.035$. Thus, relative convenience ratings were moderated by subjects' age. For those over 31, relative convenience ratings favoring the randomized strategy declined linearly as the probability of being searched increased (mean convenience = 1.17, 0.89, and 0.49, for the search conditions $p = 1/10$, 1/4, and 1/2, respectively). For those under 31, relative convenience ratings demonstrated no monotonic relationship with the probability of being searched (mean relative convenience = 0.63, 0.36, and 0.63, for the search conditions $p = 1/10$, 1/4, and 1/2, respectively). Neither main effects for sex nor age nor the sex by age interaction was significant; similarly, the sex by search condition interaction was also not significant.

4. DISCUSSION

Traditional search methods are costly and inefficient compared to randomized search strategies. Attempts to provide 100% security for all targets are practically infeasible, and almost certainly result in predictable patterns that are subject to surveillance by adversaries and are vulnerable to exploitation by intelligent, adaptive adversaries. Randomized search

strategies have the potential to provide effective security at a fraction of the cost of traditional security methods and curtail the possibility of exploiting predictable security patterns. However, public support for a search policy depends heavily on the perceived benefits to public safety. Participants in this study rated randomized security schedules as less safe than the traditional approach of searching everyone. This finding implies that—contrary to the logical and practical advantages of randomized security schedules—public perceptions of safety are not enlightened by theorems from game theory. The orthodoxy of searching everyone is preferred, even if the method is more cumbersome and theoretically less safe.

Public equity concerns and perceptions of fairness are also a clear challenge for randomized security schedules. Some have argued that any search process that allows some not to be searched is unfair on its face.[10] Proponents of randomized security schedules have argued that in fact all are treated equally, since all patrons have the same probability of being selected for a particular level of scrutiny. However, public acceptance depends not on actual fairness but rather on perceived fairness, which was directly addressed in this study. The randomized search strategy was rated as less fair than the traditional search approach across all three probability levels. One should note that the perceived unfairness of randomized searches found in this study is likely a lower bound. The randomized security schedule was described in abstract terms, but in practice, the public would observe the process of selection on a concrete, case-by-case basis. Previous psychological research on the misperception of randomness suggests that individuals will see patterns in even a perfectly randomized process.[13,14]

Most security contexts involve a tradeoff between safety and convenience. Higher levels of security may result in greater safety, but reduced convenience; similarly, minimal or no security may be very convenient for the public, but at the expense of compromised safety. It is not surprising that randomized security strategies were rated as more convenient than traditional approaches, since some fractions (1/2, 3/4, or 9/10) are allowed to bypass the search process. Moreover, the lower the proportion of subjects searched in the random process, the greater the perceived convenience of the randomized strategy compared to traditional search. Nevertheless, traditional security schedules were rated as safer than randomized schedules.

Historical data and systematic test procedures can be used to compare randomized and traditional approaches in terms of probability of detection. However, these comparisons do not capture the deterrence value of one search strategy versus another. One of the biggest challenges for evaluating search strategies is to estimate how many attacks were not attempted due to the security methods employed. This is a familiar problem associated with all counterfactuals: How do we know how many attacks would have been attempted had there been a different security method in place, or no security? Although deterrence is an important attribute for setting security policy, it is notoriously difficult to estimate accurately. Our participants rated both the traditional and randomized search strategies equivalent in deterrence value, even among the three probability conditions. Similarly, the effectiveness of the traditional and randomized search procedures was equivalent across all three probability conditions.

There were substantive differences between the traditional and random search procedures on the dimensions of safety, fairness, and convenience. Surely, an enlightened public policy would take into account how the public weights these different factors. For instance, how much convenience would the public trade off for an increase in fairness or even safety? Nevertheless, the fact that respondents were evenly divided (51% vs. 49%) in preference for the randomized versus traditional security suggests that, on average, no single factor was sufficiently salient to affect preferences. In reality, there is likely to be a great deal of heterogeneity among the public regarding its weighting of the various factors. Some might consider fairness paramount whereas for others safety might predominate. The results imply that when considering security policy, policymakers should attempt to strike a balance that does not inordinately favor any particular attribute.

This is the first experimental study to examine the public's perceptions of security schedules. Any preliminary inquiry is bound to have limitations. There is more research to be done on this issue, such as examining individual differences and different contexts in which security measures are deployed. We note that the randomized search approach utilized in this experiment is not exactly the same as that which is currently utilized at LAX. The deployed ARMOR software randomizes security for different terminals and entry points into LAX.[4,5] Individual passengers are all required to subject to individual screening. The randomized security procedure

applied in this study is applied for individuals entering an airport, stadium venue, or passing through a highway checkpoint. Traditional search methods now require searching every single person entering the airport, venue, or checkpoint. The randomized strategy we employ is akin to the procedure Transportation Security Administration (TSA) employs at airports to search carry-on bags by hand for a random sample of passengers; each passenger has an equal chance of being selected. We emphasize that the results should be accepted cautiously until replicated with different stimuli and different samples of participants. These caveats notwithstanding, we suggest that any debate about public policy for security search strategies should take into account public perceptions.

ACKNOWLEDGMENTS

This research was supported by a grant from the Army Research Office Multidisciplinary University Research Initiative (ARO MURI) number P59733-NS-MUR and from the Department of Homeland Security, Science and Technology, University Programs, number 2010ST061RE000103.

REFERENCES

1. Lewis CW. The clash between security and liberty in the U.S. response to terror. Public Administration Review, 2005; 65(1):18–30.
2. Woods J. The 911 effect: Toward a social science of terrorist threat. Social Science Journal, 2011; 48(1):213–233.
3. Homeland Security. The National Strategy for Homeland Security, 2007. Available at: http://www.whitehouse.gov/homeland/book, Accessed November 1, 2012.
4. Pita J, Jain M, Marecki J, Ordóñez F, Portway C, Tambe M, Kraus S. Deployed ARMOR protection: The application of a game theoretic model for security at the Los Angeles International Airport. In Proceedings of the 7th International Conference on Autonomous Agents and Multiagent Systems, 2008.
5. Jain M, Tsai J, Pita J, Kiekintveld C, Rathi S, Tambe M, Ordóñez F. Software assistants for randomized patrol planning for the LAX Airport Police and the Federal Air Marshal service. Interfaces, 2010; 40(4):267–290.
6. Gigerenzer G. Dread risk, September 11, and fatal traffic accidents. Psychological Science, 2004; 15(4):286–287.
7. Gigerenzer G. Out of the frying pan into the fire: Behavioral reactions to terrorist attacks. Risk Analysis, 2006; 26(2):347–351.
8. Su JC, Tran AGTT, Wirtz JG, Langteau RA, Rothman AJ. Driving under the influence (of stress): Evidence of a regional increase in impaired driving and traffic fatalities after the September 11 terrorist attacks. Psychological Science, 2009; 20(1):59–65.
9. Gaissmaier W, Gigerenzer G. 9/11, Act II: A fine-grained analysis of regional variations in traffic fatalities in the aftermath of the terrorist attacks. Psychological Science, 2012; 23(12):1449–1454.
10. Viscusi WJ, Zechhauser RJ. Sacrificing civil liberties to reduce terrorism risks. Journal of Risk and Uncertainty, 2003; 26(2/3):99–120.
11. Abelson RP. Statistics as Principled Argument. Mahwah, NJ: Eribaum Associates, 1995.
12. Tversky A, Kahneman D. Belief in the law of small numbers. Psychological Bulletin, 1971; 76(2):105–110.
13. Lopes LL. Doing the impossible: A note on induction and the experience of randomness. Journal of Experimental Psychology: Learning, Memory, and Cognition, 1982; 8(6):626–636.
14. Lopes LL, Oden GC. Distinguishing between random and nonrandom events. Journal of Experimental Psychology: Learning, Memory, and Cognition, 1987; 13(3):392–400.
15. Nickerson RS. Confirmation bias: A ubiquitous phenomenon in many guises. Review of General Psychology, 1998; 2(2):175–220.
16. Berinksy A, Huber G, Lenz G. Using Mechanical Turk as a subject recruitment tool for experimental research. Submitted for review.
17. Buhrmester M, Kwang T, Gosling SD. Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality data? Perspectives on Psychological Science, 2011; 6(1):3–5.
18. Mason W, Suri S. Conducting behavioral research on Amazon's Mechanical Turk. Behavioral Research Methods, 2012; 44(1):1–23.
19. Paolacci G, Chandler J, Ipeirotis PG. Running experiments on Amazon Mechanical Turk. Judgment and Decision Making, 2010; 5(5):411–419.
20. Oppenheimer DM, Meyvis T, Davidenko N. Instructional manipulation checks: Detecting satisficing to increase statistical power. Journal of Experimental Psychology, 2009; 45(4):776–872.
21. Gigerenzer G, Hoffrage U. How to improve Bayesian reasoning without instruction: Frequency formats. Psychological Review, 1995; 102(4):684–702.
22. Slovic P, Monahan J, MacGregor D. Violence-risk assessment and risk communication: The effects of using actual cases, providing instruction, and employing probability versus frequency formats. Law and Human Behavior, 2000; 24(3):271–296.